

University of Groningen

## Individual differences in lexical access among cochlear implant users

Nagels, Leanne; Bastiaanse, Roelien; Başkent, Deniz; Wagner, Anita

*Published in:*  
Journal of Speech Language and Hearing Research

*DOI:*  
[10.1044/2019\\_JSLHR-19-00192](https://doi.org/10.1044/2019_JSLHR-19-00192)

**IMPORTANT NOTE:** You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
2020

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Nagels, L., Bastiaanse, R., Başkent, D., & Wagner, A. (2020). Individual differences in lexical access among cochlear implant users. *Journal of Speech Language and Hearing Research*, 63(1), 286-304. [https://doi.org/10.1044/2019\\_JSLHR-19-00192](https://doi.org/10.1044/2019_JSLHR-19-00192)

**Copyright**

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

**Take-down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*

## Research Article

# Individual Differences in Lexical Access Among Cochlear Implant Users

Leanne Nagels,<sup>a,b</sup> Roelien Bastiaanse,<sup>b,c</sup> Deniz Başkent,<sup>a,d</sup> and Anita Wagner<sup>a,d</sup>

**Purpose:** The current study investigates how individual differences in cochlear implant (CI) users' sensitivity to word–nonword differences, reflecting lexical uncertainty, relate to their reliance on sentential context for lexical access in processing continuous speech.

**Method:** Fifteen CI users and 14 normal-hearing (NH) controls participated in an auditory lexical decision task (Experiment 1) and a visual-world paradigm task (Experiment 2). Experiment 1 tested participants' reliance on lexical statistics, and Experiment 2 studied how sentential context affects the time course and patterns of lexical competition leading to lexical access.

**Results:** In Experiment 1, CI users had lower accuracy scores and longer reaction times than NH listeners, particularly for nonwords. In Experiment 2, CI users' lexical competition patterns were, on average, similar to those of NH listeners, but the patterns of individual CI users varied greatly. Individual CI users' word–nonword sensitivity

(Experiment 1) explained differences in the reliance on sentential context to resolve lexical competition, whereas clinical speech perception scores explained competition with phonologically related words.

**Conclusions:** The general analysis of CI users' lexical competition patterns showed merely quantitative differences with NH listeners in the time course of lexical competition, but our additional analysis revealed more qualitative differences in CI users' strategies to process speech. Individuals' word–nonword sensitivity explained different parts of individual variability than clinical speech perception scores. These results stress, particularly for heterogeneous clinical populations such as CI users, the importance of investigating individual differences in addition to group averages, as they can be informative for clinical rehabilitation.

**Supplemental Material:** <https://doi.org/10.23641/asha.11368106>

During auditory word recognition, acoustic signals are rapidly mapped to mental representations. This process involves competition between words in the lexicon that are related to the heard sequence of sounds (e.g., Marslen-Wilson & Tyler, 1980; McClelland & Elman, 1986; Norris, 1994). Lexical competition relies on the statistical and semantic relations of words within the lexicon: the lexical statistics of the phonological form, such

as the similarity of words to other words within the lexicon (e.g., neighborhood density or frequency of occurrence) or semantic associations of words presented in the context of related words (Dahan & Tanenhaus, 2004). When speech is presented in quiet without any distortion to the signal, listeners with normal hearing (NH) resolve the initial stages of lexical access swiftly and automatically (Zhang & Samuel, 2018). In adverse conditions, when the acoustic information is degraded, listeners appear to rely more on context cues to support their interpretation of speech and, hence, on semantic associations within the lexicon (Ishida, Samuel, & Arai, 2016; Kalikow, Stevens, & Elliott, 1977; Mattys, Davis, Bradlow, & Scott, 2012; Mattys & Wiget, 2011; Pichora-Fuller, 2008). Among listeners who are exposed to degraded speech on a regular basis, such as cochlear implant (CI) users, we see great individual variability in speech perception outcomes. This variability may be partially related to different strategies to process speech, for instance, to relying more on semantic associations or on lexical statistics for lexical access as compensation for higher uncertainty about their interpretation of the speech signal, that is, lexical uncertainty. Studying the variability

<sup>a</sup>Department of Otorhinolaryngology—Head & Neck Surgery, University Medical Center Groningen, University of Groningen, the Netherlands

<sup>b</sup>Center for Language and Cognition Groningen, University of Groningen, the Netherlands

<sup>c</sup>National Research University Higher School of Economics, Moscow, Russia

<sup>d</sup>Research School of Behavioural and Cognitive Neurosciences, Graduate School of Medical Sciences, University of Groningen, the Netherlands

Correspondence to Leanne Nagels: [leanne.nagels@rug.nl](mailto:leanne.nagels@rug.nl)

Editor-in-Chief: Frederick (Erick) Gallun

Editor: Alexander L. Francis

Received December 21, 2018

Revision received May 24, 2019

Accepted September 18, 2019

[https://doi.org/10.1044/2019\\_JSLHR-19-00192](https://doi.org/10.1044/2019_JSLHR-19-00192)

**Disclosure:** The authors have declared that no competing interests existed at the time of publication.

in the strategies of processing speech among CI users can give us insight into how individual listeners adapt to degraded speech, and it can also contribute to the further development of more individualized hearing rehabilitation and diagnostics. In this study, we conducted two experiments to investigate individual differences in CI users' strategies to process speech and how these are affected by lexical uncertainty. In Experiment 1, we investigated lexical uncertainty by studying participants' ability to distinguish words from nonwords when processing single words. In Experiment 2, we investigated how CI users' reliance on sentential context when accessing the lexicon during the processing of continuous speech relates to this measure of lexical uncertainty (Experiment 1) and their clinical word recognition scores.

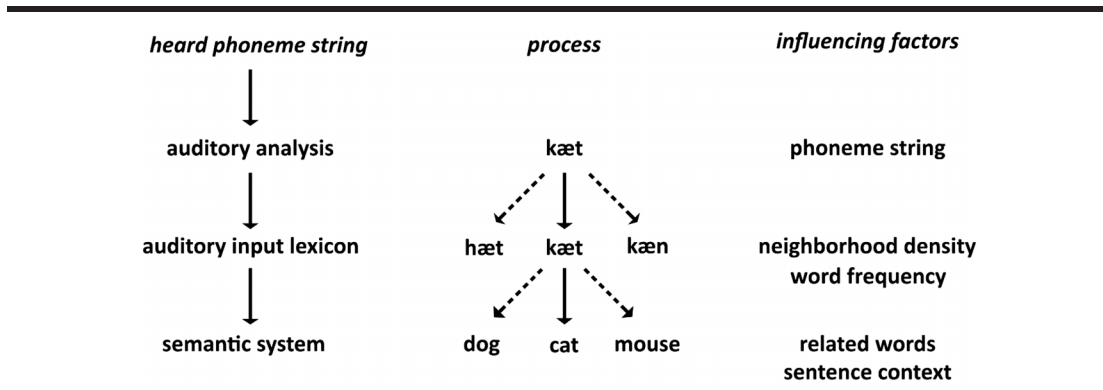
Long-term exposure to degraded speech may alter lexical access, as it can lead to coarser phonological representations (Lyxell et al., 1998) or increased lexical uncertainty (McMurray, Farris-Trimble, Seedorff, & Rigler, 2016), or it can change the manner in which phonological and lexical cues are processed (Mattys et al., 2012). However, there is great individual variability in the manner by which and the degree to which listeners adapt to degraded speech (Ishida et al., 2016), which is especially relevant for hearing-impaired populations that are faced with degraded speech on a daily basis, such as CI users (Pisoni, Kronenberger, Harris, & Moberly, 2018). A CI is a prosthetic hearing device that enables individuals with profound sensorineural hearing loss to partially regain hearing and speech perception. The speech signal perceived via the device is spectrotemporally degraded due to a combination of factors related to signal processing, the nerve electrode interface, and etiology (see a review by Başkent, Gaudrain, Tamati, & Wagner, 2016). The degraded signal makes speech processing generally more demanding and challenging for CI users (Stickney, Zeng, Litovsky, & Assmann, 2004; Vitevitch, Pisoni, Kirk, Hay-McCutcheon, & Yount, 2000; Wagner, Nagels, Toffanin, Opie, & Başkent, 2019; Winn, 2016). Several hearing and device-related factors contribute to individual variability, such as the duration of experience with the CI or the duration and age of onset of severe-to-profound hearing loss (Blamey et al., 2013; Lazard et al., 2012). Commonly, clinical assessment of CI users' speech perception outcomes is based on the recognition of single words, often monosyllabic and high-frequency words, and little is known about the postimplantation changes to mechanisms that underlie the processing of speech. In the Netherlands, the most commonly used test is the open-set word recognition test from the Nederlandse Vereniging voor Audiologie (NVA; Bosman & Smoorenburg, 1995), which consists of 45 lists of 12 high-frequency consonant–vowel–consonant (CVC) words. At the moment, there are no available clinical tests that can inform us about how mechanisms underlying speech processing are affected by long-term exposure and about individuals' strategies to adapt to the processing of degraded speech (Pisoni et al., 2018).

Most current theories and models of word recognition are based on observations of how NH listeners process lexical statistics and context cues in optimal conditions.

The process of understanding single words for NH listeners in ideal listening conditions can, simplified, be summarized in a so-called “word comprehension model,” such as the model sketched here (see Figure 1) that is based on the work of Ellis and Young (1996). When a phoneme string is perceived, phonemes are recognized, and the order is coded (see Figure 1: auditory analysis). When this string of phonemes is a word, the word form is recognized in the auditory input lexicon (see Figure 1), where these words are organized on the basis of their phonological structure. Words that have great phonological resemblance are associated; they are phonological neighbors. At this stage, the listener can decide whether they know a word or not without necessarily accessing the meaning. High lexical uncertainty, due to, for instance, noisy environments or speech signal degradations, can complicate this decision process, and listeners may fill in phonemes that were actually not processed via the auditory input lexicon. When a word in the auditory input lexicon is activated, its phonological neighbors are co-activated. In a normally functioning processing system, most activation is given to the target word (the word that was perceived), and the neighbors are inhibited. The target word “wins.” Neighborhood density, which is defined as the number of existing words that can be created via substitution, deletion, or addition of one phoneme (Luce, 1986), plays a role in the timing of word recognition. Competitive effects among words also depend on the characteristics of the word in a language. While for Dutch and English, it was found that the more neighbors a word has, the more time it takes to recognize that word (Luce & Pisoni, 1998; Vitevitch, Luce, Pisoni, & Auer, 1999), the opposite pattern was found for Spanish, a language with many relatively long words of mostly two or three syllables (Vitevitch & Rodríguez, 2005). Another factor that influences the activation of an entry in the auditory input lexicon is the frequency of the word; words that are highly frequent are recognized quicker than words that are less frequent (e.g., Dahan, Magnuson, & Tanenhaus, 2001; Luce, 1986; Marslen-Wilson, 1987; Taft & Hambly, 1986). A word that has been recognized in the auditory input lexicon activates the word's meaning in the semantic system (see Figure 1), where words are also stored based on semantic relations. For instance, the phonological word form /kæt/ activates the meaning of *cat*, but it also co-activates the semantically related words *dog*, *mouse*, and *pet*. When the listener accurately perceives and recognizes the phonological form, the meaning of *cat* wins, and the word is recognized.

While listening to speech, there are interactions between the levels of activation of words at various stages of analysis (e.g., Marslen-Wilson & Tyler, 1980). During the early stages of speech comprehension, phonologically related word forms compete, just as in the model described above (see Figure 1). Lexicality and the statistical probabilities of words within the auditory input lexicon, such as word frequency and neighborhood density, can facilitate or inhibit lexical selection (Luce & Pisoni, 1998; Magnuson, Dixon, Tanenhaus, & Aslin, 2007). In addition, listeners use the associative relations between words to make predictions

**Figure 1.** Schematic process of the stages of word comprehension and the factors that influence the speed of processing. The unbroken lines show activation, and the broken lines show co-activation flowed by inhibition.



about upcoming words, which can be described as priming. Models of speech perception describe this as pre-activation of a set of candidates. When listeners make predictions based on the semantic context, they can access the meaning of a word without the full auditory analysis. Furthermore, listeners can use semantic context to support or reflect on their interpretation of the speech signal, for instance, to compensate for a degraded speech signal (Winn, 2016). In ideal listening conditions, NH listeners use both sources of information efficiently (Magnuson et al., 2007).

It is yet unclear whether the supportive effects of context cues are similar when the speech signal is degraded, which is the case for CI users. Differences in linguistic processing and cognitive abilities, such as working memory capacities (Holden et al., 2013; Pisoni, 2000), reweighting of segmental cues (Moberly et al., 2014), or the effects of lexical statistics of words (Vitevitch et al., 2000), as well as lexical uncertainty can also contribute to individual variability among CI users. For instance, research with CI users has reported individual differences in the effects of lexical statistics, such as neighborhood density and phonotactic probability, based on their speech perception abilities (Vitevitch et al., 2000), or task demands, such as open-set versus closed-set word recognition (Sommers, Kirk, & Pisoni, 1997). A classic psycholinguistic paradigm to study the structure of the mental lexicon and the effect of lexical statistics on timing and accuracy of lexical access is auditory lexical decision (Goldinger, 1996). This paradigm has been used to study priming effects (e.g., Neely, Keefe, & Ross, 1989) and the structure of the mental lexicon and lexical access in different clinical populations, such as aphasic patients, children with specific language impairment, or CI users (e.g., Blumstein, Milberg, Dworetzky, Rosen, & Gershberg, 1991; Edwards & Lahey, 1996; Kirk, Pisoni, & Osberger, 1995; Vitevitch et al., 2000). Vitevitch et al. (2000) used an auditory lexical decision task in which CI users with high and low word recognition scores had to indicate whether the stimulus was an existing word, for example, *boat*, or a nonword, for example, *sep*. They found that CI users with high speech perception scores responded faster to words than nonwords, which resembles the response

pattern that is commonly found in NH listeners (Chambers & Forster, 1975). CI users with low speech perception scores responded equally fast to nonwords as to words and, thus, did not show an effect of lexicality, which may be caused by coarser phonological representations or may be a listener's strategy to compensate for degradation in the signal and lexical uncertainty by filling in information (Başkent, Clarke, et al., 2016; Bhargava, Gaudrain, & Başkent, 2016; Warren, 1970; Winn, 2016).

Eye tracking has been widely used to study lexical competition processes and real-time speech processing, primarily by using the visual-world paradigm (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). Dahan and Tanenhaus (2004) showed that the provision of constraining thematic constraints given by the main verb in Dutch constrained the number of lexical candidates that were considered by NH listeners during lexical competition. They used a visual-world paradigm that displayed images of a target (e.g., *baby* [baby]); a semantic competitor (e.g., *worm* [worm]), which was a possible continuation of the main verb that was used (e.g., *kruipen* [to crawl]); a phonological competitor (e.g., *beker* [cup]); and an unrelated distractor (e.g., *hortensia* [hydrangea]). The main verb preceded the target in the context condition and, thus, provided thematic constraints (e.g., *Vandaag kruipt de baby een stuk verder*. [Today the baby crawls a bit further.]). In the neutral condition, however, the target was preceded by a neutral auxiliary or modal verb that did not provide any thematic constraints (e.g., *Vandaag is de baby een stuk verder gekropen*. [Today the baby has crawled a bit further.]). The proportion of gaze fixations toward the phonological competitor in the context condition did not differ from those toward the unrelated distractor. The phonological competitor was thus not considered as a potential lexical candidate when the preceding verb provided thematic constraints. While supportive context information reduces lexical competition between phonological competitors in NH listeners in ideal listening conditions, this appears to differ in degraded speech conditions (Farris-Trimble, McMurray, Cigrand, & Tomblin, 2014; Wagner, Pals, de Blecourt, Sarampalis, & Başkent, 2016). Degraded speech has been demonstrated to



increase listeners' uncertainty about their interpretation of speech (Mattys et al., 2012), which changes the lexical competition patterns in NH listeners (Wagner et al., 2016). Wagner et al. (2016) found that, when NH listeners were presented with simulated CI (vocoded) speech, the integration of verb-based thematic constraints was delayed, and listeners' fixations toward the semantic competitor were strongly reduced. However, although degradation seems to limit the ability of the NH listeners to use context information to confine lexical candidates, it does not follow that this is also true for CI users, who may have developed compensation strategies to deal with degraded speech.

As the speech signal that is received via the CI is spectrotemporally degraded, CI users may not be able to process and integrate thematic constraints in a timely manner to confine lexical candidates. However, Huang, Newman, Catalano, and Goupell (2017) showed that CI users were able to use supportive prosodic cues to constrain the number of lexical candidates in a visual-world paradigm task, whereas NH listeners presented with simulated CI speech did not. This poses an interesting question about what specific cues CI users may use to compensate for degraded speech, as, for instance, CI users show a weaker perception of prosody (Chatterjee & Peng, 2008; Peng, Chatterjee, & Lu, 2012), although they can still utilize voice pitch (Fuller et al., 2014; Gaudrain & Başkent, 2018). CI users may thus process supportive context information differently from listeners with NH presented with simulated CI speech. Such a discrepancy between CI users and NH listeners presented with CI simulated speech was also found for phonemic restoration (Bhargava et al., 2016). Phonemic restoration is an auditory continuity illusion in which listeners are unaware of disruptions, such as a cough or noise, due to the restoration of phonemes based on top-down information (Samuel, 1981). While it is possible that the acoustic simulations of CIs do not entirely replicate the reduced acoustic-phonetic cues available to actual CI users, another possibility is that CI users may learn to make better use of degraded cues due to long-term exposure to degraded speech and experience with the CI (Başkent, Gaudrain, et al., 2016; Rouger et al., 2007).

Farris-Trimble et al. (2014) used a visual-world paradigm task to study the time course of lexical competition in postlingually deaf CI users, who were deafened and implanted after the completion of language development. Relative to NH listeners, lexical competition in CI users was generally delayed, as shown by a delay in and a lower proportion of fixations toward the target (e.g., *wizard*) as well as a higher proportion of fixations toward phonologically related competitors (e.g., *whistle* and *lizard*) but not toward the unrelated distractor (e.g., *baggage*). Differences in lexical competition patterns between NH listeners and CI users were thus primarily of a quantitative nature. CI users' proportion and timing of fixations toward the competitors differed from those of NH listeners, but their lexical competition patterns were similar. McMurray, Farris-Trimble, and Rigler (2017) demonstrated that prelingually deaf CI users, who were congenitally deaf or became deaf early (before the age of 3 years in that study),

demonstrated a wait-and-see strategy and showed different lexical competition patterns than NH listeners. The prelingually deaf CI users looked less at the target (e.g., *wizard*) and the phonological competitor (e.g., *whistle*) and more toward the rhyme competitor (e.g., *lizard*) than NH listeners, unlike postlingually deaf CI users (Farris-Trimble et al., 2014). Thus, differences in lexical competition were of a more qualitative nature, as prelingually deafened CI users differed from NH listeners not only in the proportion and timing of fixations toward competitors but also in their patterns of lexical competition.

### *The Current Study*

As most lexical access research has focused on how lexical statistics and context information are generally processed by NH listeners in ideal listening conditions, where listeners can make use of all sources of information, not much is known about individual differences in lexical access. This issue is especially relevant for listening in adverse conditions and for clinical populations. When CI users are treated as a group, lexical access is assumed to be delayed (Farris-Trimble et al., 2014; Huang et al., 2017; McMurray et al., 2017), but a large amount of individual variability in CI users' speech perception abilities makes it difficult to treat this population as a homogeneous group. Previous research has primarily tried to explain individual variability among CI users based on hearing and device-related factors (Blamey et al., 2013; Lazard et al., 2012), such as the amount of experience with the CI and differences in cognitive abilities such as working memory capacities or linguistic processing (Vitevitch et al., 2000). Here, we will focus on the linguistic mechanisms that are involved, such as listeners' use of lexical statistics and reliance on sentential context.

In this study, we conducted two experiments to investigate individual differences in CI users' strategies to process speech and how these are affected by lexical uncertainty. We conducted two experiments to study differences in the effects of lexical statistics on the processing of single words (Experiment 1) and differences in the reliance on context information on the processing of continuous speech (Experiment 2). For our first experiment, we used an auditory lexical decision task to investigate the effects of lexical statistics on the accuracy and timing of lexical access. In our second experiment, we studied individual CI users' real-time processing of sentential context and how it relates to their clinical speech perception scores and word-nonword sensitivity (as captured in Experiment 1). We used a visual-world paradigm task that replicated the earlier described experiments on verb-based thematic constraints (Dahan & Tanenhaus, 2004; Wagner et al., 2016). We hypothesized that CI users would demonstrate lower sensitivity to word-nonword differences and higher reliance on sentential context cues, such as verb-based thematic constraints, than NH listeners, due to increased lexical uncertainty about their interpretation of the spectrotemporally degraded speech signal. CI users were expected to show individual differences

in lexical uncertainty and, hence, differences in the correct categorization of nonwords. Particularly nonwords with high neighborhood densities and frequencies were expected to be difficult to categorize, as these lexical features make nonwords more word-like, which has been demonstrated to increase a lexical response bias (Luce & Pisoni, 1998). Moreover, we expected CI users with relatively low word–nonword sensitivity (Experiment 1), due to higher lexical uncertainty, to also, as a consequence, show higher reliance on sentential context information to resolve lexical competition (Experiment 2), as a compensation strategy. Finally, we expected that CI users' word–nonword sensitivity, reflecting lexical uncertainty, would explain a different part of the variability among CI users' strategies to process continuous speech than their clinical word recognition scores, reflecting speech audiometric thresholds, as previous research found discrepancies between both abilities in CI users (Amichetti, Atagi, Kong, & Wingfield, 2018; Winn, 2016).

## Experiment 1: Processing of Lexical Statistics

### Method

#### Participants

Fifteen CI users and 14 NH listeners participated in the study. All participants were native speakers of Dutch, were right-handed, and reported no speech or language disorders. The demographic characteristics of both participant groups are summarized in Table 1. CI users were recruited through advertisements and via the clinic of the Department of Otorhinolaryngology, University Medical Center Groningen. The NH control group was matched in age (range of  $\pm 5$  years) and gender with the CI group.

All CI participants were postlingually deaf and had a minimum of 2 years of experience with the implant. The individual characteristics of CI users are included in Appendix. The hearing thresholds of NH listeners were measured via a short audiometric test before data collection. NH listeners were only included if their hearing thresholds were below 25 dB HL measured at audiometric frequencies

from 500 to 4000 kHz. This is a more tolerant norm for normal hearing that takes minimal age-related hearing-level changes into account, used in analogy to previous studies (Saija, Akyürek, Andringa, & Başkent, 2014). All participants were given detailed information about the study before participation, and they signed a written informed consent form prior to data collection. Ethical approval of the study was given by the Medical Ethical Committee of the University Medical Center Groningen.

### Materials

We constructed an auditory lexical decision task with 50 words, for example, *weken* [weeks], and 50 nonwords, for example, *saren*, as well as four practice items, all of which were balanced for syllable length (one or two syllables), log frequency (range: 0.06–5.02), and phonological neighborhood density (range: 0–28). The frequency of the stimuli was determined via the SUBTLEX-NL database (Keuleers, Brysbaert, & New, 2010). For nonwords, the mean frequency of neighboring words was used. Although the frequency of occurrence of nonwords is technically 0, during lexical decision, they do compete with existing neighboring words that are activated. Previous research has demonstrated that the mean frequency of neighboring words of nonwords, “neighborhood frequency” (Luce & Pisoni, 1998) or “mean log-frequency weighted neighborhood density” (Vitevitch & Luce, 1998), has significant effects on the accuracy and reaction times of NH listeners for categorizing both words and nonwords. Therefore, for nonwords, we computed a functional mean frequency of all phonological neighboring words of the nonword, such as *taken*, *sokken*, and *smaken*, from the Dutch CLEARPOND (Cross-Linguistic Easy-Access Resource for Phonological and Orthographic Neighborhood Densities) database (Marian, Bartolotti, Chabal, & Shook, 2012). Phonological lexical neighbors were defined as existing words that can be created via substitution, deletion, or addition of one phoneme (Luce, 1986). The 50 nonwords were derived from existing words from similar frequency and phonological neighborhood density cohorts as the word stimuli and then by substituting, deleting, or adding one

**Table 1.** Demographic characteristics of participant groups.

Participant group	Variable	<i>M</i> ( <i>SD</i> )	Range
CI users	Age (years)	56.31 (14.58)	30–73
	Education (Verhage scale <sup>a</sup> )	5.47 (0.74)	5–7
	Experience with CI (years)	5.27 (3.51)	2–13
	Age at CI implantation (years)	52.87 (14.77)	28–72
	Age (years)	55.63 (11.02)	25–72
NH listeners	Education (Verhage scale <sup>a</sup> )	6.07 (0.73)	5–7

Note. CI = cochlear implant; NH = normal-hearing.

<sup>a</sup>Participants' education level is defined according to the classification based on the Dutch education system by Verhage (1964), ranging from 1 (*only primary education*) to 7 (*university-level education*). The education level of our participants was 5 (MBO; *middelbaar beroepsonderwijs*), 6 (HBO; *hoger beroepsonderwijs*), or 7 (WO; *wetenschappelijk onderwijs*), respectively.

phoneme to turn them into nonwords. For instance, the existing word *maken* [to make] was turned into the nonword *saken* by substituting the /m/ with an /s/. An overview of the stimuli with their respective lexical statistics can be found in Supplemental Material S1.

The stimuli were recorded in an anechoic room at a sampling rate of 44 kHz, spoken by a female native speaker of Dutch with a standard Dutch accent. The average fundamental frequency of the speaker was 250 Hz, and the average duration of the stimuli was 634 ms, ranging from 356 to 807 ms. The presentation level of the stimuli was equalized to a root-mean-square level of 65 dB SPL by using Praat software (Boersma & Weenink, 2018).

### Procedure and Setup

Before the experiment, an audiometric screening was done with NH listeners to ensure that their hearing thresholds agreed with our criteria for NH. Subsequently, four practice trials were presented to the participants to familiarize them with the task. Participants were seated in a soundproof booth at a distance of approximately 50–60 cm from a 17-in. LCD computer screen. The participants were instructed to press, as fast as possible, a keyboard key marked with a red sticker on the left side of the keyboard when they heard an existing word and a keyboard key marked with a yellow sticker on the right side of the keyboard when they heard a nonexisting word. Accuracy scores and reaction times from stimulus offset until participants' keypress were measured by using MATLAB (MathWorks Inc., 2012). Before the stimulus presentation, a black cross was displayed at the center of the screen for 500 ms to draw participants' attention to the center of the screen. Subsequently, stimuli were presented to participants via an AudioFire4 sound card (Echo Digital Audio) and played on a Tannoy Precision 8D speaker behind the computer monitor. This study was part of a larger project in which, in addition to measuring accuracy scores and reaction times, participants' pupillary responses were measured via an EyeLink II eye tracker during the task to investigate individual differences in listening effort. These pupillometry data are presented and discussed in Wagner et al. (2019). Here, the only minor consequence of collecting pupillometry data was that each trial was preceded by a screen during which participants were instructed to blink in order to decrease the amount of blinking during stimulus presentation. The total duration of the experiment was approximately 10 min.

### Data Analysis

To measure the effects of lexicality (word vs. nonword), frequency, and phonological neighborhood density on lexical decision, accuracy scores and the log-transformed reaction times of participants were collected and analyzed using the lme4 package (Bates et al., 2014) in R (R Core Team, 2013). Frequency values were log-transformed and centered, and phonological neighborhood density values were centered for data analysis, considering the less skewed distributions.

Accuracy scores were modeled using a mixed-effects logistic regression model with three-way interactions between Group (NH vs. CI), Stimulus Type (word vs. nonword), and Neighborhood Density (centered); Group, Stimulus Type, and Frequency (log-transformed and centered); and random intercepts for participants and items, in lme4 syntax:  $\text{accuracy} \sim \text{Group} * \text{Stimulus Type} * \text{Neighborhood Density} + \text{Group} * \text{Stimulus Type} * \text{Frequency} + (1 | \text{participant}) + (1 | \text{item})$ . This enables us to account for the individual variability in the data that is introduced by differences in the intercepts for individual participants and items. Model comparison was done using backward stepwise selection in which individual fixed effects were individually removed from the full model to determine their contribution to the model fit using the likelihood-ratio test. The change in model fit as a consequence of removing a fixed effect was measured using analysis of variance (ANOVA) chi-square tests. Reaction times of trials that received correct responses were log-transformed and modeled using a linear mixed-effects model with the same three-way interactions as for the accuracy model and random intercepts for participants and items, in lme4 syntax:  $\text{reaction time} \sim \text{Group} * \text{Stimulus Type} * \text{Neighborhood Density} + \text{Group} * \text{Stimulus Type} * \text{Frequency} + (1 | \text{participant}) + (1 | \text{item})$ .

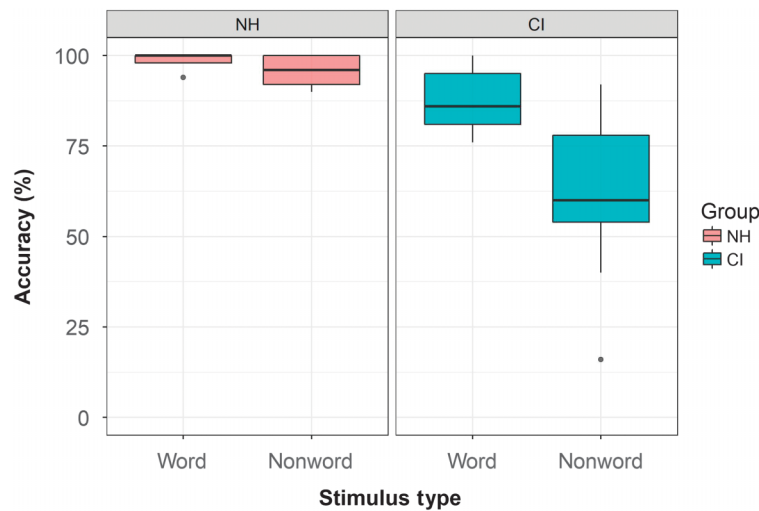
## Results

### Accuracy Scores

Figure 2 displays the median accuracy scores of NH listeners (left panel) and CI users (right panel), shown for words and nonwords in each panel. The average accuracy scores of CI users (mean words: 87.3%; mean nonwords: 62.5%) were generally lower relative to those of NH listeners (mean words: 99.0%; mean nonwords: 95.7%), but it is important to note that NH listeners approached ceiling-level performance.

A full model for accuracy scores was estimated with three-way interactions between Group (NH vs. CI), Stimulus Type (word vs. nonword), and Neighborhood Density (centered), as well as with Frequency (log-transformed and centered) as fixed effects and random intercepts for participants and items. First of all, model comparison using the ANOVA chi-square test showed that the full model with random intercepts for participants and items had a significantly better fit than the full models with only random intercepts for participants,  $\chi^2(1) = 43.12, p < .001$ , or items,  $\chi^2(1) = 77.99, p < .001$ . Backward stepwise model comparison, starting from the full model, demonstrated that the model with a two-way interaction between Stimulus Type and Frequency as well as fixed effects of Group and Neighborhood Density, in lme4 syntax:  $\text{accuracy} \sim \text{Stimulus Type} * \text{Frequency} + \text{Group} + \text{Neighborhood Density} + (1 | \text{participant}) + (1 | \text{item})$ , was the best fitting and most parsimonious model. An overview of the coefficients of the best fitting model for accuracy can be found in Supplemental Material S2. In addition, Table 2 gives an overview of the mean accuracy scores and reaction times per per Group,

**Figure 2.** Accuracy scores per group and per stimulus type. The boxplots show the median accuracy scores and the lower and upper quartiles in percentages for normal-hearing (NH) listeners (left panel in orange) and cochlear implant (CI) users (right panel in blue), and for words (left side in each panel) and nonwords (right side in each panel).



Stimulus Type, Neighborhood Density Class, and Frequency Class. Accuracy scores and reaction times were averaged for stimuli with below- and above-average neighborhood density and frequency values for this purpose. The significant effect of Group (Estimate =  $-2.90$ ,  $SE = 0.34$ ,  $p < .001$ ) indicates that CI users' accuracy scores were lower compared to those of NH listeners. Furthermore, the significant main effect of Stimulus Type (Estimate =  $-1.71$ ,  $SE = 0.22$ ,  $p < .001$ ) shows that accuracy scores for nonwords were generally lower than those for words. The significant interaction between Stimulus Type and Frequency (Estimate =  $-0.45$ ,  $SE = 0.21$ ,  $p < .05$ ) indicates that high-frequency words received higher accuracy scores than low-frequency words and that high-frequency nonwords received lower accuracy scores than low-frequency nonwords. Finally, the significant effect of Neighborhood Density (Estimate =  $-0.23$ ,  $SE = 0.11$ ,  $p < .05$ ) demonstrates that words and nonwords with high neighborhood densities had lower accuracy scores and were thus more difficult to correctly

categorize than words and nonwords with low neighborhood densities.

### Sensitivity to Word–Nonword Differences

In addition to accuracy scores, participants' sensitivity to word–nonword differences, as measured in  $d'$ , was calculated to take a potential response bias into account, which may play a role in explaining the large difference in CI users' accuracy scores for words as opposed to nonwords. Values for  $d'$  were calculated by computing the  $z$  score for the proportion of word trials in which the participant correctly categorized the stimulus as a word (hits) and subtracting the  $z$  score for the proportion of nonword trials in which participants miscategorized the stimulus as a word (false alarms). Participants'  $d'$  values are shown in Figure 3 as a function of the proportion of hits and a function of their clinical CVC scores as measured by the NVA word recognition test (Bosman & Smoorenburg, 1995).

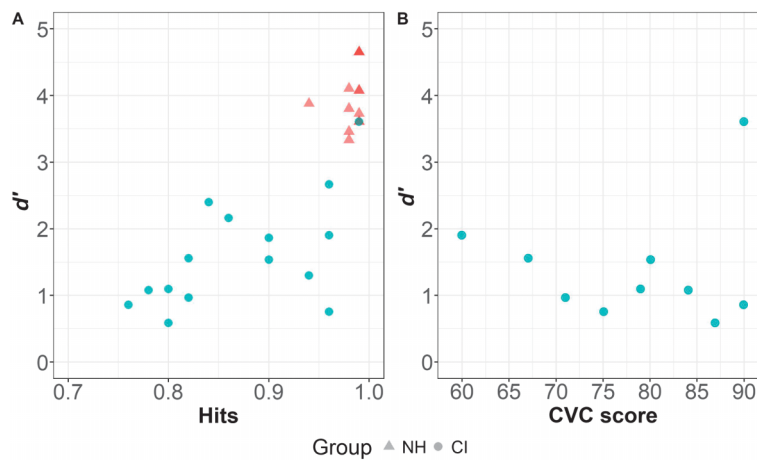
**Table 2.** Mean accuracy scores and reaction times per Group, Stimulus Type, Neighborhood Density (ND), Class (low vs. high), and Frequency Class (low vs. high).

Stimulus Type	ND Class	Frequency Class	NH accuracy (M; SD)	NH RT (M; SD)	CI accuracy (M; SD)	CI RT (M; SD)
Word	High	High	98.7%; 0.11	615 ms; 0.35	90.9%; 0.29	895 ms; 0.58
Word	High	Low	98.9%; 0.10	663 ms; 0.26	83.1%; 0.28	1001 ms; 0.56
Word	Low	High	98.9%; 0.10	542 ms; 0.33	91.8%; 0.38	760 ms; 0.58
Word	Low	Low	99.5%; 0.07	597 ms; 0.33	84.1%; 0.37	855 ms; 0.55
Nonword	High	High	95.5%; 0.21	900 ms; 0.53	47.3%; 0.50	1476 ms; 0.78
Nonword	High	Low	95.1%; 0.15	951 ms; 0.45	59.0%; 0.48	1450 ms; 0.86
Nonword	Low	High	97.8%; 0.22	789 ms; 0.62	65.6%; 0.49	1569 ms; 0.84
Nonword	Low	Low	94.5%; 0.23	849 ms; 0.48	75.9%; 0.43	1462 ms; 0.73

Note. NH = normal-hearing; RT = reaction time; CI = cochlear implant.



**Figure 3.** Individual participants'  $d'$  values for both participant groups as a function of the proportion of hits (A) and cochlear implant (CI) users'  $d'$  values as a function of their clinical consonant–vowel–consonant (CVC) scores as measured by the Nederlandse Vereniging voor Audiologie word recognition test (B). NH = normal hearing.



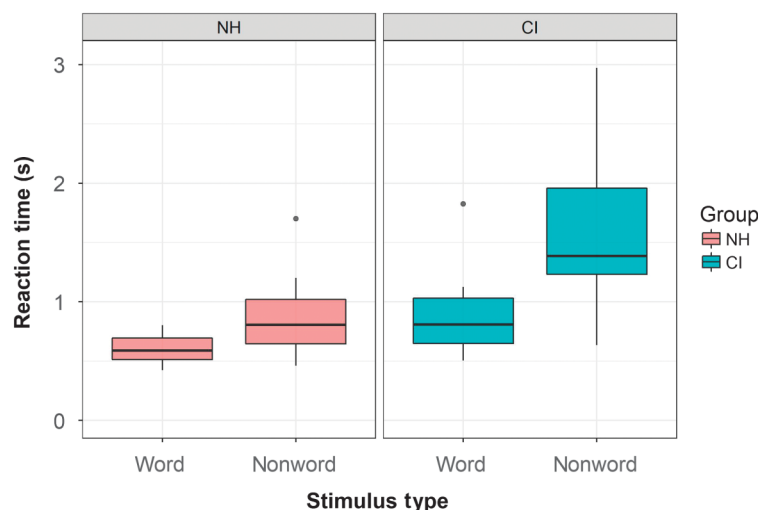
### Reaction Times

Reaction times shorter than 200 ms or longer than 3  $SDs$  above the individual participant's mean reaction time were excluded from the analysis, and only trials in which participants gave a correct response were included. This procedure removed, on average, 3.6% of the trials of NH listeners and 3.5% of the trials of CI users. For the statistical analysis, reaction times were log-transformed. Figure 4 shows the median reaction times in seconds of NH listeners (left panel) and CI users (right panel), for words and nonwords in each panel. The reaction times of CI users (mean words: 864 ms; mean non-words: 1,527 ms) were generally longer in comparison to

those of NH listeners (mean words: 589 ms; mean non-words: 886 ms).

A full model for log-transformed reaction times was estimated with three-way interactions between group (NH vs. CI), stimulus type (word vs. nonword), and neighborhood density (centered), as well as with group, stimulus type, and frequency (centered and log-transformed) as fixed effects and random intercepts for participants and items. Model comparison using the ANOVA chi-square test showed that the full model with random intercepts for participants and items had a significantly better fit than the full models with only random intercepts for participants,  $\chi^2(1) = 52.40$ ,  $p < .001$ , or items,  $\chi^2(1) = 698.68$ ,  $p < .001$ . According to

**Figure 4.** Reaction times per group and per stimulus type. The boxplots show the median reaction times and the lower and upper quartiles in seconds for listeners with normal-hearing (NH) listeners (left panel in orange) and cochlear implant (CI) users (right panel in blue), and for words (left side in each panel) and nonwords (right side in each panel).



backward stepwise model comparison, starting from the full model, the model with the interaction between Group and Stimulus Type, as well as with Neighborhood Density and frequency as fixed effects, had the best fitting, parsimonious model fit, in lme4 syntax: reaction time ~ Group \* Stimulus Type + Neighborhood Density + Frequency + (1|participant) + (1|item). An overview of the coefficients of the best fitting model for reaction times can be found in Supplemental Material S3. The significant main effects of Group (Estimate = 0.32,  $SE = 0.10$ ,  $p < .01$ ) and Stimulus Type (Estimate = 0.35,  $SE = 0.03$ ,  $p < .001$ ) show that CI users had longer reaction times than NH listeners and that participants generally had longer reaction times for words relative to nonwords. In addition, the significant interaction between Group and Stimulus Type (Estimate = 0.23,  $SE = 0.04$ ,  $p < .001$ ) shows that CI users particularly had significantly longer reaction times than NH listeners for nonwords compared to words. Furthermore, the significant effect of Neighborhood Density (Estimate = 0.07,  $SE = 0.01$ ,  $p < .001$ ) indicates that words and nonwords with high neighborhood densities received slower responses. Finally, the significant effect of Frequency (Estimate =  $-0.05$ ,  $SE = 0.01$ ,  $p < .001$ ) demonstrates that words and nonwords with high frequencies were responded to faster.

## Discussion

Our results demonstrate that CI users, as expected, generally had lower accuracy scores than NH listeners and that participants had lower accuracy scores for nonwords relative to words. For reaction times, the significant main effects and interaction between group and stimulus type showed that CI users had longer reaction times than NH listeners, but particularly for nonwords as opposed to words. Furthermore, responses to high-neighborhood-density stimuli were less accurate and slower, and responses to low-neighborhood-density stimuli were more accurate and faster for all participants. Finally, responses to high-frequency stimuli were faster than responses to low-frequency stimuli for all participants.

As commonly found in NH listeners (Chambers & Forster, 1975), CI users also had more difficulty correctly categorizing nonwords as opposed to words. However, this effect of lexicality—the discrepancy between accuracy and reaction times for words as opposed to nonwords—seems to be relatively enlarged for CI users. The recognition of nonwords requires listeners to inhibit the tendency to assume lexicality, which may be difficult for CI users due to higher lexical uncertainty caused by the unreliable degraded speech input signal. For instance, the nonword *dwagen* may be restored and misinterpreted as the neighboring existing word *dagen* [days] or *wagen* [wagon].

This explanation is also supported by the effect of neighborhood density on accuracy and reaction times. According to continuous mapping models, all words that partially correspond to the acoustic input signal are activated during word recognition. Hence, the number of activated lexical candidates during lexical competition is higher for high-neighborhood-density words and nonwords than for

low-neighborhood-density words and nonwords. The increased number of lexical candidates makes it more difficult to suppress a lexical response bias, which is necessary for the identification of nonwords. This is in agreement with earlier results from Luce and Pisoni (1998) and Vitevitch et al. (1999), which demonstrated that high-neighborhood-density words were more difficult to categorize than low-neighborhood-density words. Our results also show that participants had higher accuracy scores for high-frequency words than low-frequency words and lower accuracy scores for high-frequency nonwords than low-frequency nonwords. These opposing effects of frequency on word and nonword recognition have also been commonly found in previous research (Luce & Pisoni, 1998; Vitevitch & Luce, 1998). High frequency may make words easier to correctly categorize as words, but it may make it more difficult to correctly categorize nonwords. Furthermore, our results show that participants responded slower to low-frequency words and nonwords. This provides evidence for the hypothesis that, also for CI users, high-frequency words are processed and accessed faster due to lower activation thresholds, which is in line with earlier research (Dahan et al., 2001). However, it is important to take into account that we used a functional mean frequency of all phonological neighboring words for nonwords, as they technically have a frequency of 0, which may also contribute to the differential effects of frequency on words and nonwords that we found.

Finally, as is demonstrated in Figure 3, there is a substantial amount of individual variability among CI users. Some CI users performed almost as well as NH listeners, whereas other CI users scored considerably lower compared to NH listeners. These results emphasize the importance of investigating individual participants' data, particularly for clinical populations, as it may provide us with information about which compensation strategies, such as higher reliance on context information, may be more successful for the processing of continuous speech (Amichetti et al., 2018; Winn, 2016). Hence, we conducted a second experiment to further investigate how individual CI users' sensitivity to word-nonword differences, taken as a measure of CI users' lexical uncertainty, and clinical CVC word recognition scores, reflecting CI users' speech audiometric thresholds, are related to their use of sentential context when accessing the lexicon during the real-time processing of continuous speech.

---

## Experiment 2: Processing of Sentential Context Cues

### Method

#### Participants

The same group of 15 CI users and 14 NH listeners who took part in the first experiment also participated in the second experiment. Participants performed both experiments in the same test session after taking a short break.

## Materials

The materials consisted of 44 noun sets of picturable nouns, the noun sets used by Dahan and Tanenhaus (2004) and one additional and similarly constructed noun set by Wagner et al. (2016). Each noun set consisted of drawings of a target noun, a phonological competitor, a semantic competitor, and an unrelated distractor (see Figure 5). The noun sets were matched to a verb that provided thematic constraints, which were coherent with the target noun and the semantic competitor. For instance, the target noun *baby* [baby] and the semantic competitor *worm* [worm] are both semantically plausible continuations of the verb *kruipen* [to crawl]. Furthermore, a phonological competitor with the same phonological onset as the target noun was used, for example, *beker* [cup] for the target noun *baby* [baby]. Finally, an unrelated distractor was selected that was phonologically unrelated to the target noun and not semantically compatible with the verb-based thematic constraints, for example, the unrelated distractor *hortensia* [hydrangea].

The experiment was divided into two experimental blocks that each contained 22 target and 26 filler noun sets. Each experimental block was preceded by four practice trials with filler noun sets. The 60 filler noun sets consisted of a target noun, a semantic competitor, and two distractors that were phonologically unrelated to the target and not semantically compatible with the verb-based thematic constraints. In 20 filler items, the two distractors had overlapping phonological onsets to prevent participants from expecting that one of two phonologically overlapping words was always the target. The remaining 40 filler items had unrelated distractors with different phonological onsets.

Each sentence started with a short adverbial phrase, for instance, *Nog nooit* [Never before] or *Vanochtend* [This morning]. In the neutral condition, the target noun was preceded by a neutral auxiliary or modal verb (e.g., *Vandaag is de baby een stuk verder gekropen*. [Today the baby has crawled a bit further.]). In the context condition, the target

noun was preceded by the main verb (e.g., *Vandaag kruipt de baby een stuk verder*. [Today the baby crawls a bit further.]), and thematic constraints were provided before the target noun was produced.

All sentence stimuli were recorded by a male Dutch native speaker with a standard Dutch accent in an anechoic room at a sampling rate of 44 kHz. The average fundamental frequency of the speaker was 152 Hz, and the average duration of the stimuli was 2.98 s and ranged from 1.90 to 4.92 s. The presentation level of the stimuli was equalized to a root-mean-square level of 65 dB SPL by using Praat software (Boersma & Weenink, 2018).

## Procedure and Setup

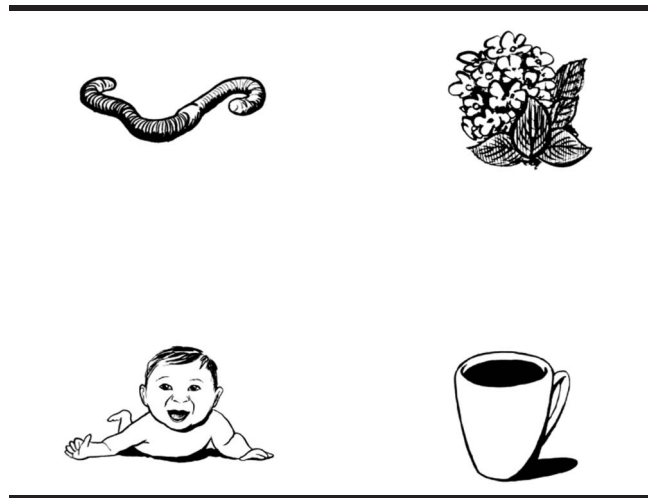
Prior to data collection, participants were familiarized with the drawings that were used in the experiment to ensure that they would correctly relate drawings to their corresponding nouns. Participants were asked to name the drawings that were presented on a computer screen and were corrected by the experimenter if necessary. During the eye-tracking experiment, participants were seated in a sound-proof booth at a distance of approximately 50–60 cm from a 17-in. LCD computer screen, where an EyeLink II eye tracker was placed on the participant's head and calibrated via a standard 9-point calibration procedure. Every five trials, a drift correction was done to ensure that the eye tracker did not lose track of the pupil. If the drift range was too large, the eye tracker was recalibrated. Eye movements were tracked by the left camera of the eye tracker using a sampling rate of 250 Hz. Gaze fixations were binned into intervals of 20 ms by taking the average of five consecutive samples. Ocular responses were recorded using the EyeLink Toolbox for MATLAB (Cornelissen, Peters, & Palmer, 2002).

At the beginning of each experimental block, four practice trials with filler noun sets were presented to the participants. Each trial was preceded by a screen during which participants were instructed to blink. After this, a red cross was presented at the center of the screen for 500 ms. Subsequently, four drawings from the respective noun set were displayed on the screen (see Figure 5), while participants heard a sentence that was presented through an AudioFire4 sound card (Echo Digital Audio) and played on a Tannoy Precision 8D speaker behind the computer monitor. Participants were instructed to click on the drawing of the target noun that was mentioned in the sentence as soon as they heard it. Each noun set was presented to the participant in either the neutral or context condition. Responses were measured and registered by using MATLAB (MathWorks Inc., 2012). The total duration of the experiment was approximately 30 min.

## Data Analysis

Gaze fixations within the time window of 200–2,000 ms after the target noun onset were analyzed as binomial responses. Trials in which an incorrect response was given or

Figure 5. Example item: drawings of a target noun set.



that contained eye blinks or artifacts that were longer than 300 ms were excluded from the analysis. Eye blinks that were shorter than 300 ms were linearly interpolated based on the median value of 50 samples preceding and following the blink.

Growth curve analysis (Mirman, 2014) was used to analyze the proportions of gaze fixations of participants toward the phonological and semantic competitors using the *lme4* package (Bates et al., 2014) in R (R Core Team, 2013). The time curves of gaze fixations toward the phonological and semantic competitors were modeled as fourth-order (quartic) orthogonal polynomials to see how context information affected lexical competition patterns between the competitors over time. These four time terms were used to model the average height of the curve (intercept), the steepness of the ramp and angle of the curve (linear term), the sharpness of the centered peak (quadratic term), and the sharpness of the additional peaks and the curvature in the tails (cubic and quartic terms) of the time curves of gaze fixations toward the competitors (Mirman, 2014). We interpret these terms as indications of the duration of the time course and the degree of lexical competition that was experienced by participants. The steepness of the ramp primarily indicates the speed at which the speech signal was processed and lexical competition was initiated. The sharpness of the centered and additional peaks as well as the curvature in the tails mainly show how fast lexical competition was resolved and the degree of certainty that participants had about their interpretation of the speech signal. Growth curve analysis enables us to look more closely at how fixation patterns develop over time and how different factors affect different parts of the general curve. The different time terms are mainly used for statistical reference, as the effects on different time terms are difficult to interpret in isolation. We used mixed-effects logistic regression models with random intercepts for participants for all time terms. Two three-way interactions between Group (NH vs. CI), Condition (neutral vs. context), and all time terms, and Condition,  $d'$  (participants' sensitivity to word–nonword differences), and all time terms were added to the model, in *lme4* syntax: fixations phonological/semantic competitor ~ (linear + quadratic + cubic + quartic) \* Group \* Condition + (linear + quadratic + cubic + quartic) \* Condition \*  $d'$  + (linear + quadratic + cubic + quartic | participant). In addition, we performed an analysis on a subset of 10 CI users from the hearing clinic of the University Medical Center Groningen, whose NVA clinical CVC word recognition scores (Bosman & Smoorenburg, 1995) we had, in *lme4* syntax: fixations phonological/semantic competitor ~ (linear + quadratic + cubic + quartic) \* Condition \* CVC scores + (linear + quadratic + cubic + quartic) \* Condition \*  $d'$  + (linear + quadratic + cubic + quartic | participant). The values for CVC scores were centered, considering the less skewed distribution. Backward stepwise model comparison was done using the likelihood-ratio test to measure the change in model fit as a consequence of removing a fixed effect. ANOVA chi-square tests were used to measure the change in model fit.

## Results

Figure 6 shows the time course of gaze fixations toward the target (green lines), the phonological competitor (red lines), the semantic competitor (magenta lines), and the unrelated distractor (black lines). Proportions and time curves of gaze fixations for NH listeners and CI users for the neutral and context conditions are averaged across participants and trials. A 95% confidence interval for gaze fixations of both participant groups toward the target and competitors is also included in the figure.

### *Gaze Fixations Toward the Phonological Competitor*

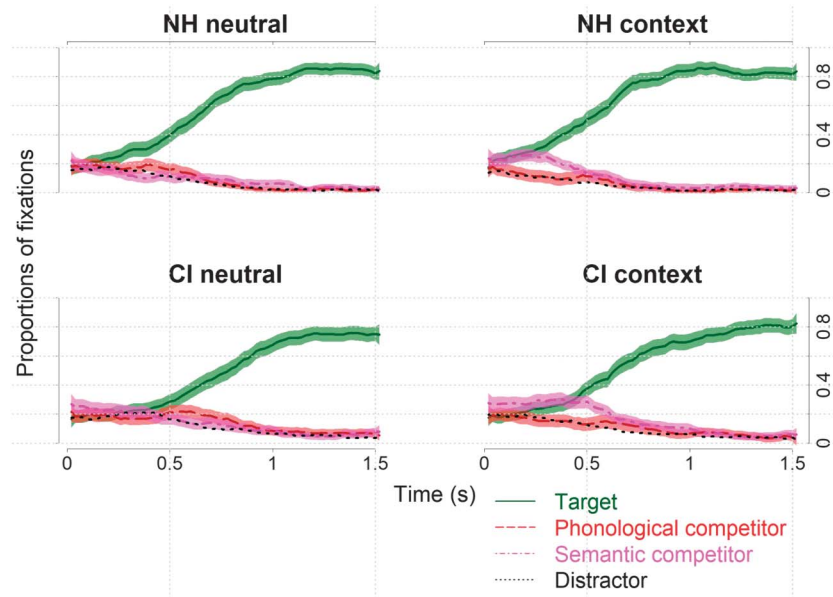
Gaze fixations toward the phonological competitor were modeled as a fourth-order (quartic) orthogonal polynomial using growth curve analysis with random intercepts for participants for all time terms. Backward stepwise model comparison showed that the full model with the three-way interaction between Group, Condition, and all time terms, and the three-way interaction between  $d'$ , Condition, and all time terms had a significantly better fit than the models without either of the three-way interactions,  $\chi^2(10) = 176.63$ ,  $p < .001$ ;  $\chi^2(10) = 197.43$ ,  $p < .001$ , or models in which an individual fixed effect was removed from the interaction. An overview of the coefficients of the best fitting model for fixations toward the phonological competitor can be found in Supplemental Material S4.

According to the model, there was a significant main effect of Condition on the fixations toward the phonological competitor (Estimate =  $-0.54$ ,  $SE = 0.03$ ,  $p < .001$ ), showing that participants, overall, looked less toward the phonological competitor in the neutral condition than the context condition. The significant interaction between Group and Condition (Estimate =  $0.25$ ,  $SE = 0.05$ ,  $p < .001$ ) indicates that CI users, overall, looked more toward the phonological competitor in the neutral condition than the context condition compared to NH listeners. The significant effects of the interaction between Group and Condition on the linear term (Estimate =  $-4.28$ ,  $SE = 1.01$ ,  $p < .001$ ), the quadratic term (Estimate =  $5.97$ ,  $SE = 1.05$ ,  $p < .001$ ), the cubic term (Estimate =  $5.82$ ,  $SE = 1.03$ ,  $p < .001$ ), and the quartic term (Estimate =  $3.43$ ,  $SE = 0.99$ ,  $p < .001$ ) show that the angle of the ramp was steeper and that the areas around the peaks were sharper for NH listeners than for CI users, particularly in the neutral condition compared to the context condition.

The significant interaction between  $d'$  and Condition (Estimate =  $0.17$ ,  $SE = 0.02$ ,  $p < .001$ ) shows that participants, overall, looked more toward the phonological competitor in the neutral condition than the context condition as their word–nonword sensitivity was lower, reflecting higher lexical uncertainty. The significant effects of the interaction between  $d'$  and Condition on the linear term (Estimate =  $1.29$ ,  $SE = 0.37$ ,  $p < .001$ ), the quadratic term (Estimate =  $2.54$ ,  $SE = 0.37$ ,  $p < .001$ ), and the cubic term (Estimate =  $2.11$ ,  $SE = 0.37$ ,  $p < .001$ ) show that the angle



**Figure 6.** Time course of gaze fixations shown per group and per condition toward the target (green solid lines), the phonological competitor (red solid lines), the semantic competitor (magenta dotted lines), and the distractor (black dotted lines) for normal-hearing (NH) listeners (upper panels) and cochlear implant (CI) users (lower panels), and for the neutral condition (left panels) and the context condition (right panels).



of the ramp was less steep and that the areas around the peaks were less sharp as participants' word-nonword sensitivity was lower, particularly in the neutral condition.

### Gaze Fixations Toward the Semantic Competitor

Gaze fixations toward the semantic competitor were modeled as a fourth-order (quartic) orthogonal polynomial using growth curve analysis with random intercepts for participants for all time terms. Backward stepwise model comparison showed that the full model with the three-way interaction between group, condition, and all time terms as the three-way interaction between Group, Condition, and all time terms, and the three-way interaction between  $d'$ , Condition, and all time terms had a significantly better fit than the models without either of the three-way interactions,  $\chi^2(10) = 307.52, p < .001$ ;  $\chi^2(10) = 350.26, p < .001$ , or models where an individual fixed effect was removed from the interaction. An overview of the coefficients of the best fitting model for fixations toward the semantic competitor can be found in Supplemental Material S5.

According to the model, there was a significant main effect of Condition on the fixations toward the semantic competitor (Estimate = 0.36,  $SE = 0.03, p < .001$ ), showing that participants, overall, looked more toward the semantic competitor in the context condition than the neutral condition. The significant interaction between Group and Condition (Estimate = -0.56,  $SE = 0.05, p < .001$ ) shows that CI users, overall, looked more toward the semantic competitor in the context condition than the normal condition relative to NH listeners. There were significant effects of the interaction between Group and Condition on the

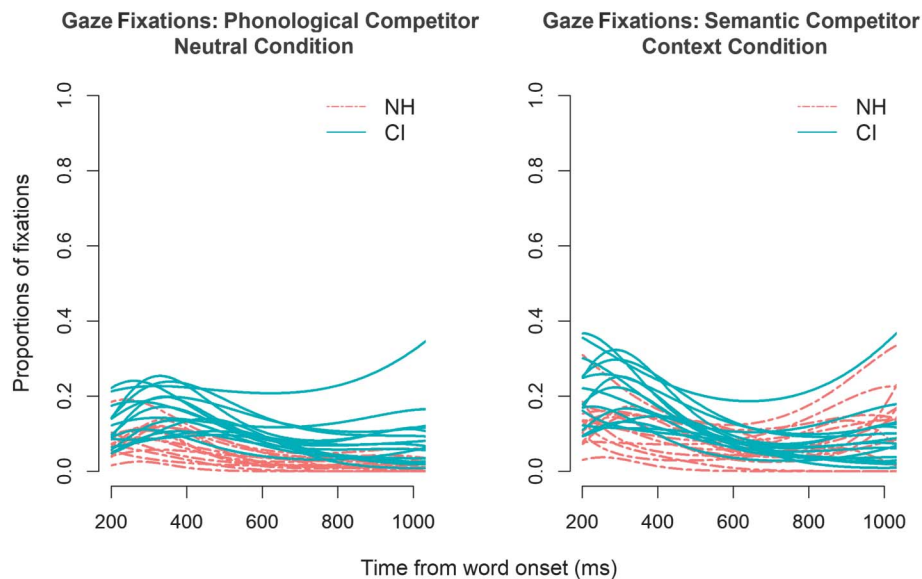
linear term (Estimate = -10.56,  $SE = 0.86, p < .001$ ), the quadratic term (Estimate = 4.29,  $SE = 0.97, p < .001$ ), the cubic term (Estimate = 6.64,  $SE = 0.92, p < .001$ ), and the quartic term (Estimate = -4.04,  $SE = 0.89, p < .001$ ). These results show that the angle of the ramp and the areas around the peaks were sharper for NH listeners than for CI users, particularly in the context condition compared to the neutral condition.

The significant interaction between  $d'$  and Condition (Estimate = -0.26,  $SE = 0.02, p < .001$ ) shows that participants with relatively low word-nonword sensitivity, overall, looked more toward the semantic competitor than participants with high word-nonword sensitivity in the context condition compared to the neutral condition. The significant effects of the interaction between  $d'$  and Condition on the linear term (Estimate = -3.04,  $SE = 0.32, p < .001$ ), the quadratic term (Estimate = 3.61,  $SE = 0.36, p < .001$ ), and the quartic term (Estimate = -1.98,  $SE = 0.33, p < .001$ ) show that the angle of the ramp and the areas around the peaks were sharper as participants' word-nonword sensitivity was lower, particularly in the context condition.

### Gaze Fixation Patterns of Individual CI Users

In addition to the general between-groups analysis, we performed a second analysis on the gaze fixation data of a subset of 10 CI users from the hearing clinic of the University Medical Center Groningen, whose NVA clinical CVC test scores we had, to look more closely at differences in the processing of context information between individual CI users. As is shown by the gaze fixation patterns in Figure 7, there seems to be considerable individual

**Figure 7.** Time course of gaze fixations toward the phonological competitor in the neutral condition (left panel) and the context condition (right panel), and for normal-hearing (NH) listeners (orange lines) and cochlear implant (CI) users (blue lines).



variability within the timing of gaze fixations of CI users, whereas for NH listeners, there are merely small differences in the proportion of fixations. Furthermore, earlier research of, among others, Amichetti et al. (2018) and Winn (2016) suggests that CI users' speech intelligibility scores do not correspond to their use of sentential context for speech perception. Hence, we wanted to investigate whether CI users' CVC test scores, a clinical measure of speech perception, and word–nonword sensitivity ( $d'$ ), taken as a measure of lexical uncertainty, both equally contributed to explaining individual variability in the real-time speech processing fluency of context information in CI users.

### Model Comparison

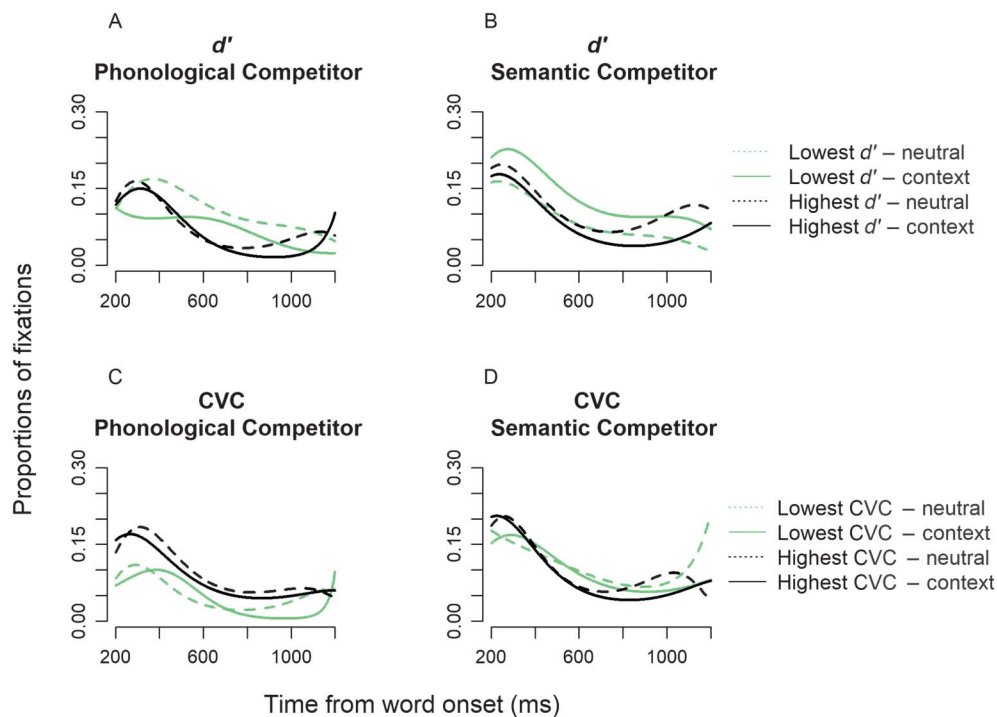
A fourth-order (quartic) orthogonal polynomial was used to model the gaze fixations toward the phonological and semantic competitors of a subset of the CI users by means of growth curve analysis (Mirman, 2014). We used mixed-effects logistic regression models with random intercepts for participants for all time terms. For the phonological competitor, backward stepwise model comparison showed that the full model with the three-way interaction between  $d'$ , Condition, and all time terms, and the three-way interaction between CVC, Condition, and all time terms had a significantly better fit than the models without either of the three-way interactions,  $\chi^2(10) = 370.58, p < .001$ ;  $\chi^2(10) = 203.41, p < .001$ , or models in which an individual fixed effect was removed from the interaction. An overview of the coefficients of the best fitting model for CI users' fixations toward the phonological competitor can be found in

Supplemental Material S6. For the semantic competitor, the model with the three-way interaction between  $d'$ , Condition, and all time terms, and the three-way interaction between CVC, Condition, and all time terms also had a significantly better fit than the models without either of the three-way interactions,  $\chi^2(10) = 102.07, p < .001$ ;  $\chi^2(10) = 272.71, p < .001$ , or models in which an individual fixed effect was removed. An overview of the coefficients of the best fitting model for CI users' fixations toward the semantic competitor can be found in Supplemental Material S7.

### Effect of Word–Nonword Sensitivity

Figures 8A and 8B show the time course of fixations toward the phonological and semantic competitors for the lowest centered  $d'$ , which was obtained by CI users, of  $-2.28$  and the highest centered  $d'$  of  $0.74$ , based on the best fitting model of our data. For the phonological competitor, there was a significant main effect of Condition (Estimate =  $-0.29$ ,  $SE = 0.04, p < .001$ ), indicating that CI users, overall, looked more toward the phonological competitor in the neutral condition than the context condition. The significant interaction between  $d'$  and Condition (Estimate =  $0.11$ ,  $SE = 0.02, p < .001$ ) shows that the lower the CI users' word–nonword sensitivity, the higher the overall proportion of fixations toward the phonological competitor in the neutral condition, indicating a higher degree of phonological competition. Furthermore, the significant interaction between  $d'$  and the quadratic term (Estimate =  $4.55$ ,  $SE = 1.86, p < .05$ ) shows that the areas around the centered peaks were shallower, and thus, the decrease in fixations toward the phonological competitor was slower, as CI users'  $d'$  was lower. Finally,

**Figure 8.** Fitted model of the time course of gaze fixations for the subgroup of cochlear implant users, for the lowest and highest centered  $d'$  values (A and B) and the lowest and highest centered consonant–vowel–consonant (CVC) scores (C and D). The solid lines show the modeled effects for the context condition, and the dashed lines show the modeled effects for the neutral condition.



there were significant effects of the interaction between  $d'$  and Condition on the linear term (Estimate =  $-1.54$ ,  $SE = 0.43$ ,  $p < .001$ ), the quadratic term (Estimate =  $1.71$ ,  $SE = 0.45$ ,  $p < .001$ ), and the cubic term (Estimate =  $3.88$ ,  $SE = 0.43$ ,  $p < .001$ ), indicating that the lower the CI users' word–nonword sensitivity, the shallower the angle of the ramp as well as areas around the peaks in the context condition, and hence, the later the occurrence of the decrease in fixations toward the phonological competitor.

For the semantic competitor, there was a significant main effect of Condition (Estimate =  $-0.18$ ,  $SE = 0.04$ ,  $p < .001$ ), showing that CI users, overall, looked less toward the semantic competitor in the context condition. Figure 8B suggests that this is mainly caused by a small effect of context on the fixation patterns of CI users with relatively high  $d'$  values. Furthermore, there was a significant interaction between  $d'$  and Condition (Estimate =  $-0.33$ ,  $SE = 0.02$ ,  $p < .001$ ), showing that the overall proportion of fixations toward the semantic competitor in the context condition became higher, as CI users' word–nonword sensitivity was lower. The significant interaction between  $d'$  and Condition also had a significant effect on the linear term (Estimate =  $-3.61$ ,  $SE = 0.41$ ,  $p < .001$ ), showing the angle of the ramp became less steep in the context condition as CI users' word–nonword sensitivity was lower. To summarize, the presence of context affected the fixation patterns of CI users with relatively low word–nonword sensitivity more than for

those of CI users with relatively high word–nonword sensitivity. In the context condition, CI users with relatively low word–nonword sensitivity looked more at the semantic competitor and less at the phonological competitor, whereas context did not have a clear effect on the fixation patterns of CI users with relatively high word–nonword sensitivity, although it resolved lexical competition faster.

### Effect of NVA CVC Scores

Figures 8C and 8D show the time course of fixations toward the phonological and semantic competitors for the lowest centered CVC score, which was obtained by CI users, of  $-18.53$  and the highest centered CVC score of  $11.47$ , based on the best fitting model of our data. For the phonological competitor, there was a significant interaction between CVC and Condition (Estimate =  $0.01$ ,  $SE = 0.002$ ,  $p < .001$ ), showing that the lower the CI score, the lower the overall proportion of fixations toward the phonological competitor in the context condition, indicating a lower degree of phonological competition. Finally, there were significant effects of the interaction between CVC and Condition on the linear term (Estimate =  $0.64$ ,  $SE = 0.05$ ,  $p < .001$ ), the quadratic term (Estimate =  $0.26$ ,  $SE = 0.04$ ,  $p < .001$ ), the cubic term (Estimate =  $-0.53$ ,  $SE = 0.04$ ,  $p < .001$ ), and the quartic term (Estimate =  $-0.29$ ,  $SE = 0.04$ ,  $p < .001$ ), showing that the lower the CI users' CVC score, the

shallower the angle of the ramp as well as areas around the peaks in the context condition.

For the semantic competitor, there was no significant interaction between CVC and Condition (Estimate =  $-0.002$ ,  $SE = 0.002$ ,  $p = .29$ ), indicating that the CI users' CVC score did not contribute to the overall proportion of fixations toward the semantic competitor. However, the interaction between CVC and Condition did have significant effects on the linear term (Estimate =  $0.11$ ,  $SE = 0.04$ ,  $p < .01$ ), the quadratic term (Estimate =  $0.24$ ,  $SE = 0.04$ ,  $p < .001$ ), the cubic term (Estimate =  $0.24$ ,  $SE = 0.14$ ,  $p < .001$ ), and the quartic term (Estimate =  $0.22$ ,  $SE = 0.13$ ,  $p < .001$ ), showing that the angle of the ramp and the areas around the curves became shallower as CI users' CVC score was lower in the context condition. To summarize, when CI users' CVC score was relatively low, the overall fixations toward the phonological competitor were lower, and the decline in fixations toward the phonological and semantic competitors was shallower and occurred later than for CI users with relatively high CVC scores. In Figure 8, we can see that CVC scores did not seem to reflect differences between the neutral and context conditions, but rather differences in the competition between phonologically similar words and their overall speech of resolving lexical competition.

## Discussion

Our results demonstrate that the lexical competition patterns of CI users, as reflected by the time course of their gaze fixations averaged over groups, were generally similar to those of NH listeners. A general effect of context information was found in both participant groups. CI users were thus, on average, able to timely integrate context information to constrain the lexical candidates that were considered. Most of the CI users and NH listeners looked less at the phonological competitor and more at the semantic competitor when context information was provided. Hence, lexical competition primarily took place between the target and the phonological competitor in the neutral condition and between the target and the semantic competitor in the context condition. This fixation pattern does differ from NH listeners who were tested in degraded speech conditions (Wagner et al., 2016). Prediction of lexical candidates can merely be achieved when the context information, in this case, thematic constraints provided by the main verb, is processed and recognized on time. CI users' experience with degraded speech may have enabled them to process the context information fast enough to constrain the number of activated lexical candidates, whereas this was not the case for NH listeners who were tested in degraded speech conditions. This benefit of exposure to degraded speech and experience with the CI is in line with earlier reports (Blamey et al., 2013; Lazard et al., 2012). Testing adults with NH in degraded speech can, in some ways, be informative of CI users' speech processing, but the results of adults with NH tested in degraded speech seem to be missing the effects of experience, adaptation,

and compensation among CI users (Farris-Trimble et al., 2014; Huang et al., 2017; Winn, 2016).

Furthermore, these findings are in agreement with continuous mapping models, such as the TRACE (McClelland & Elman, 1986) and Shortlist (Norris, 1994; Norris & McQueen, 2008) models, that describe that context information constrains the number of activated lexical candidates during lexical competition. These results also fit with earlier research that has demonstrated degraded speech conditions, and hence higher lexical uncertainty, increase listeners' reliance on supportive cues, such as sentential context cues, rather than decreasing it (Ishida et al., 2016; Mattys et al., 2012). However, although our results show that context information affects the lexical competition patterns of CI users in a similar manner as NH listeners, there are differences in their processing fluency of context information. CI users, overall, looked more at the phonological and semantic competitors compared to NH listeners, and the decrease in and peak of fixations were shallower and occurred later. CI users were thus able to timely integrate and use context information to constrain the number of lexical candidates, but the overall time course of lexical competition was still prolonged relative to NH listeners. The integration of context information can support listeners' interpretation of the speech signal, but it may not necessarily reduce the time course of lexical competition. According to Farris-Trimble et al. (2014), the delay may be caused by the harsh onset of the processed speech signal received by CI users. Another possibility is that the unreliability of the acoustic input signal increases lexical uncertainty in CI users and, therefore, increases the activation levels of competing lexical candidates, which prolongs the time course of lexical competition (Mattys et al., 2012; McMurray et al., 2016). Some CI users may also use context information as a way to affirm their interpretation of the speech signal (Winn, 2016), which may also prolong the time course of lexical competition. It would be interesting for future research to try to distinguish CI users who use context information to make predictions about the upcoming speech signal from CI users who use context information to affirm their interpretation of the speech signal. For instance, by looking at phono-semantic priming (Huang & Snedeker, 2011; Yee & Sedivy, 2006), the effects of semantic information are expected to be earlier and more rapid than in the current study.

The general analysis of gaze fixation patterns implies that the difference between NH listeners and CI users is primarily quantitative and not qualitative: We found differences merely in the proportion and duration of fixations toward the competitors, but similar fixation patterns between groups. However, the additional analysis for individual CI users shows that there is a large amount of individual variability in how similar their gaze fixation patterns are to NH listeners. CI users' word-nonword sensitivity (as captured in Experiment 1) and clinical CVC word recognition scores significantly contributed to explaining differences in the proportion of fixations toward competitors in each condition, and thus explained



different parts of individual CI users' processing fluency of context information.

First, CI users' word–nonword sensitivity, an indication of their lexical uncertainty, mainly explained differences in the reliance on context information to resolve lexical competition during the processing of continuous speech (see Figures 8A and 8B). CI users with relatively low word–nonword sensitivity showed a larger effect of context on their fixations toward the competitors than CI users with relatively high word–nonword sensitivity. CI users with relatively low word–nonword sensitivity looked more toward the semantic competitor in the context condition and more toward the phonological competitor in the neutral condition, whereas there was no clear difference as a result of context on the fixation patterns of CI users with relatively high word–nonword sensitivity. However, the overall time course of lexical competition was longer as CI users' word–nonword sensitivity became lower, and lexical uncertainty higher, indicating that higher reliance on context delayed the overall time course of lexical competition. Hence, it seems that higher lexical uncertainty in CI users led to higher reliance on context information to affirm their interpretation of the speech signal (Winn, 2016).

Second, CI users' clinical CVC scores primarily explained differences in lexical competition with the phonological competitor and did not seem to reflect differences between the neutral and context conditions (see Figures 8C and 8D). Lexical competition with the phonological competitor, overall, was decreased and prolonged for CI users with relatively low CVC scores. The slower and less steep decline in fixations toward the competitors and the smaller effect of context information indicate that context and acoustic information is processed less rapidly by CI users with relatively low CVC scores. CI users with relatively low CVC scores may thus not process the onset of the target word fast enough for it to induce competition between the target and the phonological competitor (Farris-Trimble et al., 2014). The discrepancy we found between the parts of the variability that are explained by CI users' word–nonword sensitivity and CVC scores is in line with earlier research (Amichetti et al., 2018; Winn, 2016) and suggests that these measures indeed capture different mechanisms.

## General Discussion

The aim of this study was to investigate individual differences in CI users' lexical uncertainty by studying the effects of lexical statistics on CI users' ability to distinguish words from nonwords and their reliance on sentential context during the processing of continuous speech. Our results, as obtained on a group level, align with previous findings of merely quantitative differences in the proportion and time course of fixations between NH listeners and CI users (Farris-Trimble et al., 2014; Huang et al., 2017). CI users were able to timely integrate context information to constrain the considered lexical candidates similar to NH

listeners, but the time course of lexical competition was generally prolonged in CI users relative to NH listeners. Our additional within-group analysis based on sensitivity to word–nonword differences in Experiment 1 ( $d'$ ) and clinical CVC word recognition scores, however, also showed more qualitative differences in the patterns of lexical competition among CI users. We found that CI users' sensitivity to word–nonword differences (Experiment 1) explained differences in their reliance on context information to resolve lexical competition, whereas CI users' clinical CVC word recognition scores explained differences in lexical competition with the phonological competitor and the fluency of processing acoustic information.

Unlike NH listeners who were presented with degraded speech, CI users were able to timely integrate context information to constrain the considered lexical candidates (Wagner et al., 2016), and CI users showed similar effects of lexical statistics during auditory lexical decision as NH listeners. The difference between the performance of NH listeners in degraded speech conditions and that of actual CI users hints at the huge plasticity of the perceptual system, where listeners can adapt to process highly degraded speech signals more efficiently (Blamey et al., 2013; Lazard et al., 2012). On the other hand, there is a great amount of individual variability in how listeners adapt to processing degraded speech and which cues they rely on, even in NH listeners (Ishida et al., 2016). For the population of CI users, such individual differences could explain variability in speech perception outcomes, compensatory strategies, and certainty about their interpretation of the speech signal, which can be important for clinical rehabilitation since it reflects the day-to-day challenges for these listeners. It is still not well understood how CI users develop different compensation strategies to deal with degraded signals and whether a more personalized rehabilitation postimplantation would allow listeners access to various sources of cues before their strategy to compensate fossilizes. Individual differences within the population of NH listeners are often obscured in carefully controlled experiments and with homogenous groups of listeners. It seems that extreme cases, as visible in experiments with clinical populations, bring forward the need to better understand individual differences. A stronger focus on individual differences in the processing of speech is thus also important for the gain of knowledge about the plasticity of the perceptual system *per se*.

As our sample population consisted of only 15 CI users, we have to interpret our results with caution, particularly due to a large amount of variability in CI users' speech perception abilities. In addition, all CI participants in this study were relatively well-performing CI users who were satisfied with their device, which is not fully representative of the overall population of CI users. The experimental tasks required the participants to process speech presented via a loudspeaker in the auditory modality only, without any additional visual cues. Such conditions inadvertently lead to preselecting a population of CI users who show relatively better speech comprehension abilities and excluding CI users who developed

other compensatory strategies to process continuous speech in daily life, for instance, by greater reliance on visual speech information (Rouger et al., 2007). Further research with a larger sample of CI users is required to confirm the trends in strategies of speech processing that we found. However, due to the great variability in CI users' speech perception abilities, a larger sample size would also increase the variability and may be equally flawed in identifying and making claims about general patterns among CI users.

By looking at both between-groups and within-group differences, as measured via gaze fixation patterns on a task that involves the processing of continuous speech, we can gain more information about the factors that can explain differences in the processing fluency of speech of CI users (Holden et al., 2013; Pisoni, 2000). For instance, results of Vitevitch et al.'s (2000) study showed differences in CI users' processing of words and nonwords based on their word perception scores. Discrepancies between clinical test scores and tasks that look at the underlying processing mechanisms, such as auditory lexical decision, can inform us about what factors affect CI users' speech perception performance in daily life and what compensation mechanisms they employ. Furthermore, this information provides us with a better understanding of speech processing in degraded conditions (Ishida et al., 2016; Mattys et al., 2012). As mentioned earlier, general models of speech perception are often based on how NH listeners, on average, process speech in ideal listening conditions, but our findings demonstrate that there is still a large amount of individual variability caused by the manner in which listeners adapt their processing strategies when they are more uncertain about their interpretation of the speech signal. Speech comprehension with CIs is an extreme case of how individual variability in speech processing can be overlooked, but degrees of variability are likely also present among NH listeners when tested in more realistic conditions.

## Acknowledgments

This work was supported by Marie Curie Intra-European Fellowship Grant FP7-PEOPLE-2012-IEF 332402 and a MED-EL research grant, awarded to Anita Wagner; by a Vidi grant from the Netherlands Organisation for Scientific Research, Netherlands Organisation for Health Research and Development Grant 016.093.397, and funds from the Heinsius Houbolt Foundation, awarded to Deniz Başkent; and partially by Russian Federation Government Grant 14.641.31.0004 through the Center for Language and Brain, National Research University Higher School of Economics, awarded to Roelien Bastiaanse. The study is part of the research program "Healthy Ageing and Communication" of the Department of Otorhinolaryngology, University Medical Center Groningen. We would like to thank Frans Cornelissen (University Medical Center Groningen) for providing the eye tracker for this study and the Center for Information Technology of the University of Groningen for their support and for providing access to the Peregrine high-performance computing cluster. Furthermore, we are grateful to Paolo Toffanin for his technical support and advice.

## References

- Amichetti, N. M., Atagi, E., Kong, Y. Y., & Wingfield, A. (2018). Linguistic context versus semantic competition in word recognition by younger and older adults with cochlear implants. *Ear and Hearing, 39*(1), 101–109.
- Başkent, D., Clarke, J., Pals, C., Benard, M. R., Bhargava, P., Saija, J., . . . Gaudrain, E. (2016). Cognitive compensation of speech perception with hearing impairment, cochlear implants, and aging: How and to what degree can it be achieved. *Trends in Hearing, 20*, 1–16.
- Başkent, D., Gaudrain, E., Tamati, T. N., & Wagner, A. (2016). Perception and psychoacoustics of speech in cochlear implant users. In A. T. Cacace, E. de Kleine, A. G. Holt, & P. van Dijk (Eds.), *Scientific foundations of audiology: Perspectives from physics, biology, modeling, and medicine* (pp. 285–319). San Diego, CA: Plural.
- Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., . . . Grothendieck, G. (2014). *Pack- age 'lme4.'* Vienna, Austria: R Foundation for Statistical Computing.
- Bhargava, P., Gaudrain, E., & Başkent, D. (2016). The intelligibility of interrupted speech: Cochlear implant users and normal hearing listeners. *Journal of the Association for Research in Otolaryngology, 17*(5), 475–491.
- Blamey, P., Artieres, F., Başkent, D., Bergeron, F., Beynon, A., Burke, E., . . . Lazard, D. S. (2013). Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants: An update with 2251 patients. *Audiology & Neurotology, 18*(1), 36–47.
- Blumstein, S. E., Milberg, W. P., Dworetzky, B., Rosen, A., & Gershberg, F. (1991). Syntactic priming effects in aphasia: An investigation of local syntactic dependencies. *Brain and Language, 40*(3), 393–421.
- Boersma, P., & Weenink, D. (2018). Praat: Doing phonetics by computer (Version 6.0.37) [Computer software]. Retrieved from <http://www.praat.org>
- Bosman, A. J., & Smoorenburg, G. F. (1995). Intelligibility of Dutch CVC syllables and sentences for listeners with normal hearing and with three types of hearing impairment. *Audiology, 34*(5), 260–284.
- Chambers, S. M., & Forster, K. I. (1975). Evidence for lexical access in a simultaneous matching task. *Memory & Cognition, 3*(5), 549–559.
- Chatterjee, M., & Peng, S.-C. (2008). Processing F0 with cochlear implants: Modulation frequency discrimination and speech intonation recognition. *Hearing Research, 235*(1–2), 143–156.
- Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The Eye-link Toolbox: Eye tracking with MATLAB and the Psychophysics Toolbox. *Behavior Research Methods, Instruments, & Computers, 34*(4), 613–617.
- Dahan, D., Magnuson, J. S., & Tanenhaus, M. K. (2001). Time course of frequency effects in spoken-word recognition: Evidence from eye movements. *Cognitive Psychology, 42*(4), 317–367.
- Dahan, D., & Tanenhaus, M. K. (2004). Continuous mapping from sound to meaning in spoken-language comprehension: Immediate effects of verb-based thematic constraints. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*(2), 498–513.
- Edwards, J., & Lahey, M. (1996). Auditory lexical decisions of children with specific language impairment. *Journal of Speech and Hearing Research, 39*(6), 1263–1273.

- Ellis, A. W., & Young, A. W. (1996). *Human cognitive neuropsychology: A textbook with readings*. Hove, United Kingdom: Psychology Press.
- Farris-Trimble, A., McMurray, B., Cigrand, N., & Tomblin, J. B. (2014). The process of spoken word recognition in the face of signal degradation. *Journal of Experimental Psychology: Human Perception and Performance*, 40(1), 308–327.
- Fuller, C. D., Gaudrain, E., Clarke, J. N., Galvin, J. J., Fu, Q.-J., Free, R. H., & Başkent, D. (2014). Gender categorization is abnormal in cochlear implant users. *Journal of the Association for Research in Otolaryngology*, 15(6), 1037–1048.
- Gaudrain, E., & Başkent, D. (2018). Discrimination of voice pitch and vocal-tract length in cochlear implant users. *Ear and Hearing*, 39(2), 226–237.
- Goldinger, S. D. (1996). Auditory lexical decision. *Language and Cognitive Processes*, 11(6), 559–568.
- Holden, L. K., Finley, C. C., Firszt, J. B., Holden, T. A., Brenner, C., Potts, L. G., . . . Skinner, M. W. (2013). Factors affecting open-set word recognition in adults with cochlear implants. *Ear and Hearing*, 34(3), 342–360.
- Huang, Y. T., Newman, R. S., Catalano, A., & Goupell, M. J. (2017). Using prosody to infer discourse prominence in cochlear-implant users and normal-hearing listeners. *Cognition*, 166, 184–200.
- Huang, Y. T., & Snedeker, J. (2011). Cascading activation across levels of representation in children's lexical processing. *Journal of Child Language*, 38(3), 644–661.
- Ishida, M., Samuel, A. G., & Arai, T. (2016). Some people are “More Lexical” than others. *Cognition*, 151, 68–75.
- Kalikow, D. N., Stevens, K. N., & Elliott, L. L. (1977). Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *The Journal of the Acoustical Society of America*, 61(5), 1337–1351.
- Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods*, 42(3), 643–650.
- Kirk, K. I., Pisoni, D. B., & Osberger, M. J. (1995). Lexical effects on spoken word recognition by pediatric cochlear implant users. *Ear and Hearing*, 16(5), 470–481.
- Lazard, D. S., Vincent, C., Venail, F., Van de Heyning, P., Truy, E., Sterkers, O., . . . Blamey, P. J. (2012). Pre-, per- and postoperative factors affecting performance of postlinguistically deaf adults using cochlear implants: A new conceptual model over time. *PLOS ONE*, 7(11), 1–11.
- Luce, P. A. (1986). *Neighborhoods of words in the mental lexicon. Research on speech perception* (Technical Report No. 6). Bloomington: Indiana University Press.
- Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, 19(1), 1–36.
- Lyxell, B., Andersson, J., Andersson, U., Arlinger, S., Bredberg, G., & Harder, H. (1998). Phonological representation and speech understanding with cochlear implants in deafened adults. *Scandinavian Journal of Psychology*, 39(3), 175–179.
- Magnuson, J. S., Dixon, J. A., Tanenhaus, M. K., & Aslin, R. N. (2007). The dynamics of lexical competition during spoken word recognition. *Cognitive Science*, 31(1), 133–156.
- Marian, V., Bartolotti, J., Chabal, S., & Shook, A. (2012). CLEARPOND: Cross-Linguistic Easy-Access Resource for Phonological and Orthographic Neighborhood Densities. *PLOS ONE*, 7(8), 1–11.
- Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, 25(1), 71–102.
- Marslen-Wilson, W. D., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition*, 8(1), 1–71.
- MathWorks Inc. (2012). MATLAB: The language of technical computing. Desktop tools and development environment, version 2012a. Natick, MA: Author.
- Mattys, S. L., Davis, M. H., Bradlow, A. R., & Scott, S. K. (2012). Speech recognition in adverse conditions: A review. *Language and Cognitive Processes*, 27(7–8), 953–978.
- Mattys, S. L., & Wiget, L. (2011). Effects of cognitive load on speech recognition. *Journal of Memory and Language*, 65(2), 145–160.
- McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, 18(1), 1–86.
- McMurray, B., Farris-Trimble, A., & Rigler, H. (2017). Waiting for lexical access: Cochlear implants or severely degraded input lead listeners to process speech less incrementally. *Cognition*, 169, 147–164.
- McMurray, B., Farris-Trimble, A., Seedorff, M., & Rigler, H. (2016). The effect of residual acoustic hearing and adaptation to uncertainty on speech perception in cochlear implant users: Evidence from eye-tracking. *Ear and Hearing*, 37(1), e37–e51.
- Mirman, D. (2014). *Growth curve analysis and visualization using R*. Boca Raton, FL: CRC Press.
- Moberly, A. C., Lowenstein, J. H., Tarr, E., Caldwell-Tarr, A., Welling, D. B., Shahin, A. J., & Nittrouer, S. (2014). Do adults with cochlear implants rely on different acoustic cues for phoneme perception than adults with normal hearing. *Journal of Speech, Language, and Hearing Research*, 57(2), 566–582.
- Neely, J. H., Keefe, D. E., & Ross, K. L. (1989). Semantic priming in the lexical decision task: Roles of prospective prime-generated expectancies and retrospective semantic matching. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(6), 1003–1019.
- Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, 52(3), 189–234.
- Norris, D., & McQueen, J. M. (2008). Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review*, 115(2), 357–395.
- Peng, S.-C., Chatterjee, M., & Lu, N. (2012). Acoustic cue integration in speech intonation recognition with cochlear implants. *Trends in Amplification*, 16(2), 67–82.
- Pichora-Fuller, M. K. (2008). Use of supportive context by younger and older adult listeners: Balancing bottom-up and top-down information processing. *International Journal of Audiology*, 47(Suppl. 2), S72–S82.
- Pisoni, D. B. (2000). Cognitive factors and cochlear implants: Some thoughts on perception, learning, and memory in speech perception. *Ear and Hearing*, 21(1), 70–78.
- Pisoni, D. B., Kronenberger, W. G., Harris, M. S., & Moberly, A. C. (2018). Three challenges for future research on cochlear implants. *World Journal of Otorhinolaryngology—Head and Neck Surgery*, 3(4), 240–254.
- R Core Team. (2013). *R: A language and environment for statistical computing*. Retrieved from <http://www.R-project.org/>
- Rouger, J., Lagleyre, S., Frayssé, B., Deneve, S., Deguine, O., & Barone, P. (2007). Evidence that cochlear-implanted deaf patients are better multisensory integrators. *Proceedings of the National Academy of Sciences of the United States of America*, 104(17), 7295–7300.
- Saija, J. D., Akyürek, E. G., Andringa, T. C., & Başkent, D. (2014). Perceptual restoration of degraded speech is preserved with advancing age. *Journal of the Association for Research in Otolaryngology*, 15(1), 139–148.
- Samuel, A. G. (1981). Phonemic restoration: Insights from a new methodology. *Journal of Experimental Psychology: General*, 110(4), 474–494.



- Sommers, M. S., Kirk, K. I., & Pisoni, D. B. (1997). Some considerations in evaluating spoken word recognition by normal-hearing, noise-masked normal-hearing, and cochlear implant listeners. I: The effects of response format. *Ear and Hearing*, 18(2), 89–99.
- Stickney, G. S., Zeng, F.-G., Litovsky, R., & Assmann, P. (2004). Cochlear implant speech recognition with speech maskers. *The Journal of the Acoustical Society of America*, 116(2), 1081–1091.
- Taft, M., & Hambly, G. (1986). Exploring the cohort model of spoken word recognition. *Cognition*, 22(3), 259–282.
- Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, 268(5217), 1632–1634.
- Verhage, F. (1964). *Intelligence and age: Research on Dutch people aged twelve to seventy-seven years old*. Assen, the Netherlands: Van Gorcum.
- Vitevitch, M. S., & Luce, P. A. (1998). When words compete: Levels of processing in perception of spoken words. *Psychological Science*, 9(4), 325–329.
- Vitevitch, M. S., Luce, P. A., Pisoni, D. B., & Auer, E. T. (1999). Phonotactics, neighborhood activation, and lexical access for spoken words. *Brain and Language*, 68(1–2), 306–311.
- Vitevitch, M. S., Pisoni, D. B., Kirk, K. I., Hay-McCutcheon, M., & Yount, S. L. (2000). Effects of phonotactic probabilities on the processing of spoken words and nonwords by adults with cochlear implants who were postlingually deafened. *The Volta Review*, 102(4), 283–302.
- Vitevitch, M. S., & Rodríguez, E. (2005). Neighborhood density effects in spoken word recognition in Spanish. *Journal of Multilingual Communication Disorders*, 3(1), 64–73.
- Wagner, A. E., Nagels, L., Toffanin, P., Opie, J. M., & Başkent, D. (2019). Individual variations in effort: Assessing pupillometry for the hearing impaired. *Trends in Hearing*, 23, 1–18.
- Wagner, A. E., Pals, C., de Blecourt, C. M., Sarampalis, A., & Başkent, D. (2016). Does signal degradation affect top-down processing of speech. In P. van Dijk, D. Başkent, E. Gaudrain, E. de Kleine, A. Wagner, & C. Lanting (Eds.), *Physiology, psychoacoustics and cognition in normal and impaired hearing* (Vol. 894, pp. 297–306). New York, NY: Springer.
- Warren, R. M. (1970). Perceptual restoration of missing speech sounds. *Science*, 167(3917), 392–393.
- Winn, M. B. (2016). Rapid release from listening effort resulting from semantic context, and effects of spectral degradation and cochlear implants. *Trends in Hearing*, 20, 1–17.
- Yee, E., & Sedivy, J. C. (2006). Eye movements to pictures reveal transient semantic activation during spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(1), 1–14.
- Zhang, X., & Samuel, A. G. (2018). Is speech recognition automatic? Lexical competition, but not initial lexical access, requires cognitive resources. *Journal of Memory and Language*, 100, 32–50.

## Appendix

### Individual Characteristics of Cochlear Implant Listeners

Subject	Gender	Age	Experience CI (years)	CI Brand	Etiologies	Education <sup>a</sup>	CVC	d'
1	M	38	13	Cochlear	Meningitis	7	90	3.59
2	M	68	7	Cochlear	Infection	5	90	0.85
3	M	68	12	Cochlear	Unknown	5	84	1.06
4	M	54	2	Cochlear	Menière	5	60	1.92
5	M	61	5	Cochlear	Unknown	5	79	1.09
6	F	61	3	Advanced Bionics	Unknown	5	75	0.97
7	M	70	3	Med-EL	Unknown	5	80	1.99
8	M	58	8	Advanced Bionics	Unknown	7	71	1.03
9	M	31	3	Advanced Bionics	Unknown	6	NA	1.29
10	F	71	5	Med-EL	Menière	5	NA	2.15
11	F	63	7	Advanced Bionics	Unknown	6	NA	1.85
12	F	54	2	Cochlear	Unknown	5	NA	2.36
13	M	73	6	Cochlear	Meningitis	5	67	1.67
14	F	30	2	Cochlear	Unknown	6	NA	2.54
15	M	73	2	Advanced Bionics	Unknown	5	87	0.57

Note. CVC = consonant–vowel–consonant.

<sup>a</sup>Participants' education level is defined according to the classification based on the Dutch education system by Verhage (1964), ranging from 1 (only primary education) to 7 (university-level education). The education level of our participants was 5 (MBO; *middelbaar beroepsonderwijs*), 6 (HBO; *hoger beroepsonderwijs*), or 7 (WO; *wetenschappelijk onderwijs*), respectively.